

In [1]:

```
import plotly
import cufflinks as cf
import pandas as pd
import numpy as np
```

In [2]:

```
train = pd.read_csv("train.csv")
train
```

Out[2]:

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500
...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500

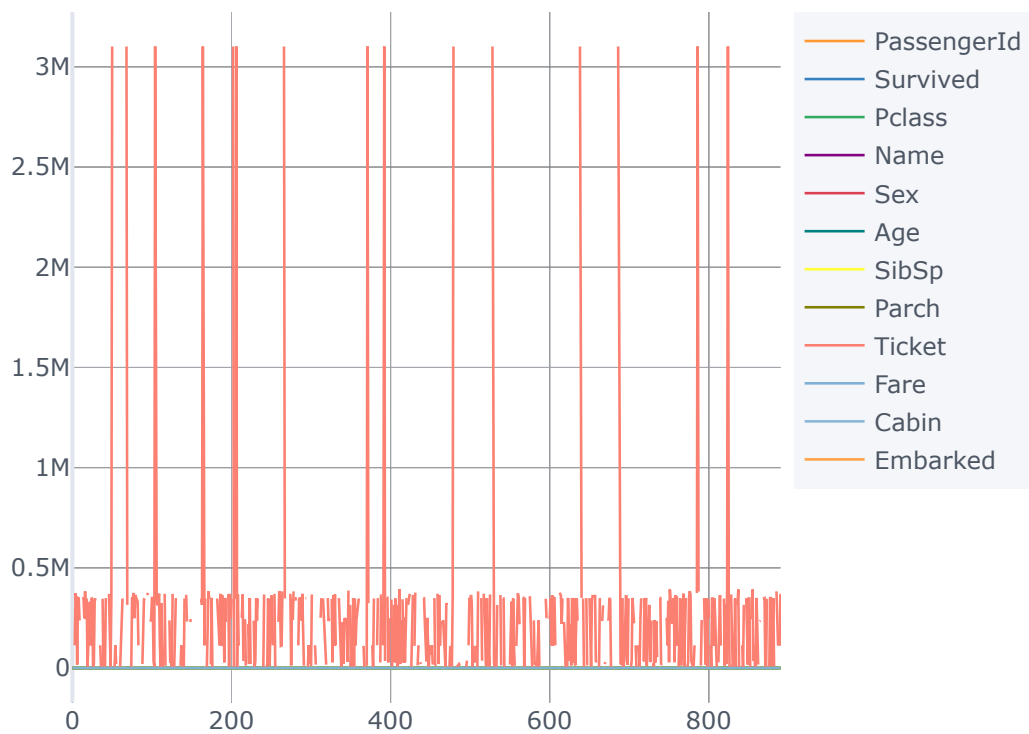
891 rows × 12 columns

In [3]:

```
# 오프라인 모드에서도 인터랙티브한 그래픽을 가능하도록 하기
# Enabling the offline mode for interactive plotting locally
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
init_notebook_mode(connected=True)
cf.go_offline()
```

In [4]:

```
train.iplot()
```

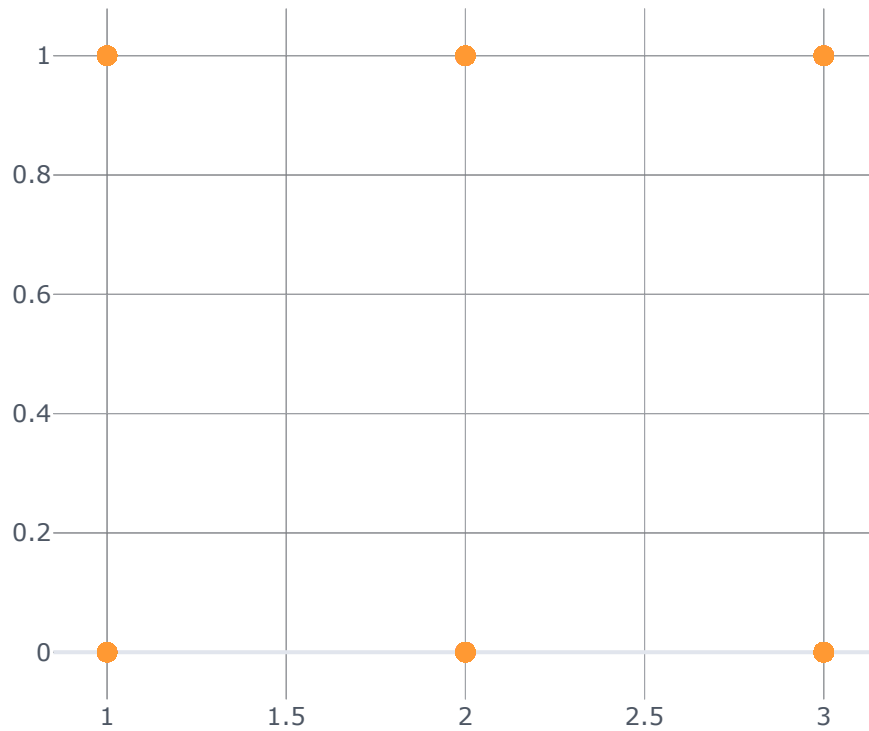
[Export to plot.ly »](#)

- 컬럼이 12개 있다.

In [5]:



```
train.ipplot(kind='scatter', x='Pclass', y='Survived', mode="markers", size=10)
```

[Export to plot.ly »](#)

- Pclass는 1,2,3의 값이 있다.
- Survived는 0,1의 값이 있다.

In [6]:



```
train.info()
```

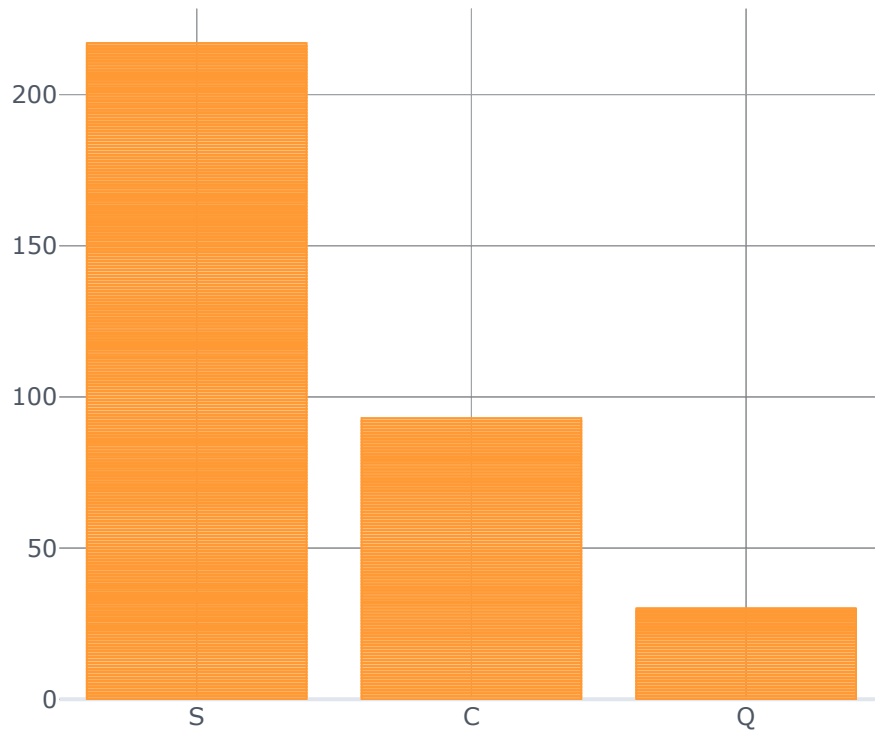
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype  
---  --
 0   PassengerId    891 non-null    int64  
 1   Survived       891 non-null    int64  
 2   Pclass         891 non-null    int64  
 3   Name           891 non-null    object  
 4   Sex            891 non-null    object  
 5   Age            714 non-null    float64 
 6   SibSp          891 non-null    int64  
 7   Parch          891 non-null    int64  
 8   Ticket         891 non-null    object  
 9   Fare           891 non-null    float64 
10   Cabin          204 non-null    object  
11   Embarked       889 non-null    object  
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

승선항 시각화

In [7]:



```
train.iplot(kind="bar", x="Embarked", y='Survived')
```

[Export to plot.ly »](#)

- 승선항 S, C, Q이 있다.
- S에서 승선한 사람이 많이 살았다.
 - 왜 S에서 승선한 사람이 많이 살았을까?

In [8]:



```
train.columns
```

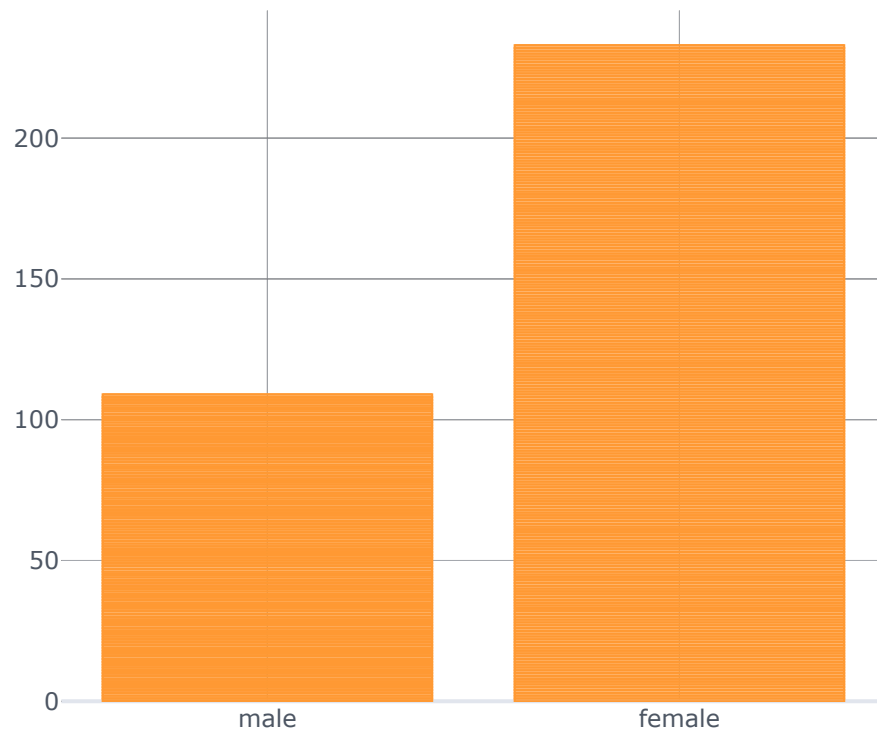
Out [8]:

```
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',  
      'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],  
      dtype='object')
```

In [9]:



```
train.iplot(kind="bar", x="Sex", y='Survived')
```

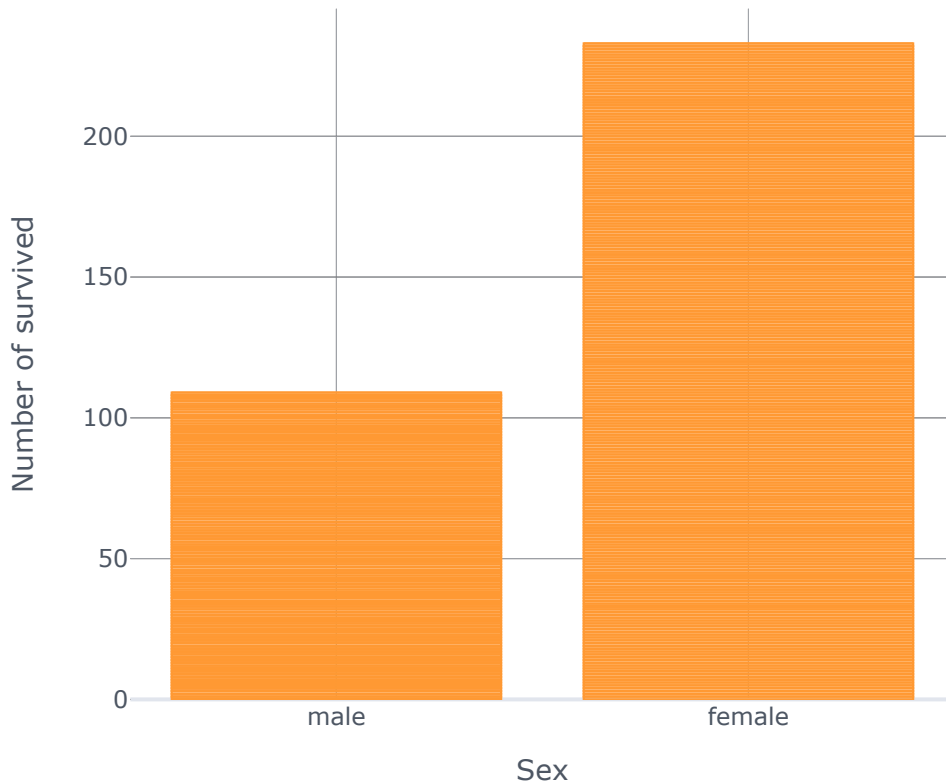
[Export to plot.ly »](#)

In [10]:

```
# x축 y축 레이블 넣기,  
train.ipplot(kind = "bar",  
              x = "Sex", y = "Survived",  
              title = "Survivors",  
              xTitle = "Sex",  
              yTitle = "Number of survived")
```



Survivors

[Export to plot.ly »](#)

- 여성이 많이 살았다.
- 2배이상의 차이가 있다.

In [11]:



```
train.columns
```

Out[11]:

```
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',  
      'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],  
      dtype='object')
```

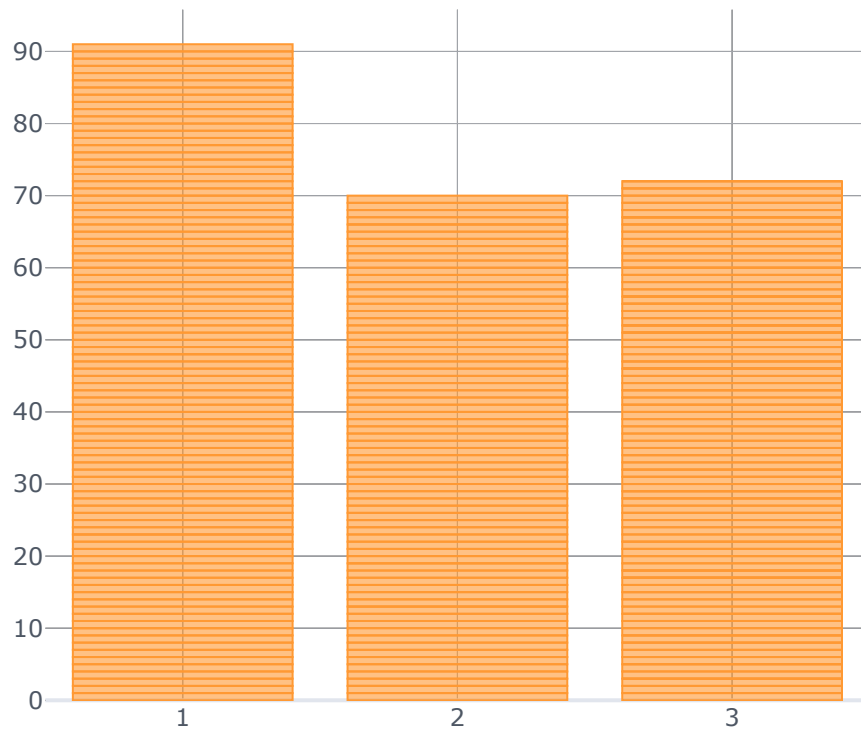
In [12]:



```
## 여성만 뽑기  
female_train = train[train.Sex=='female']
```

In [13]:

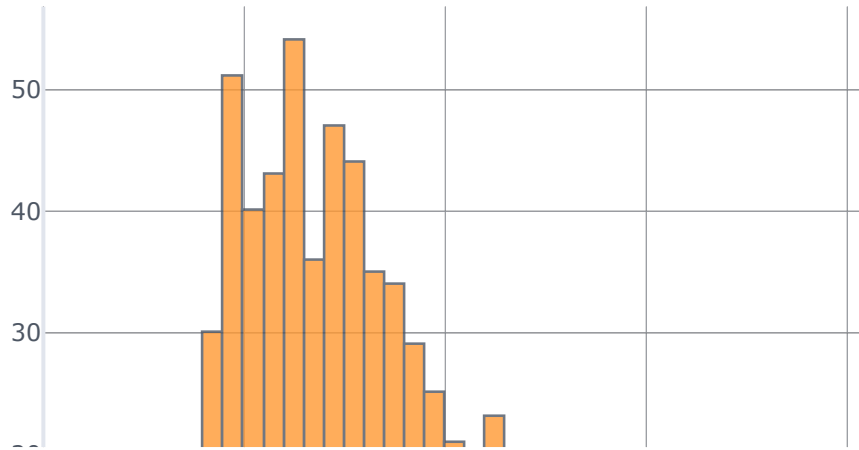
```
# 여성들 중의 pclass  
female_train.iplot(kind="bar", x="Pclass", y='Survived')
```

[Export to plot.ly »](#)

- 여성들 중의 1등급이 좀 더 많이 살아남았다.

In [14]:

```
train['Age'].ipplot(kind='hist')
```

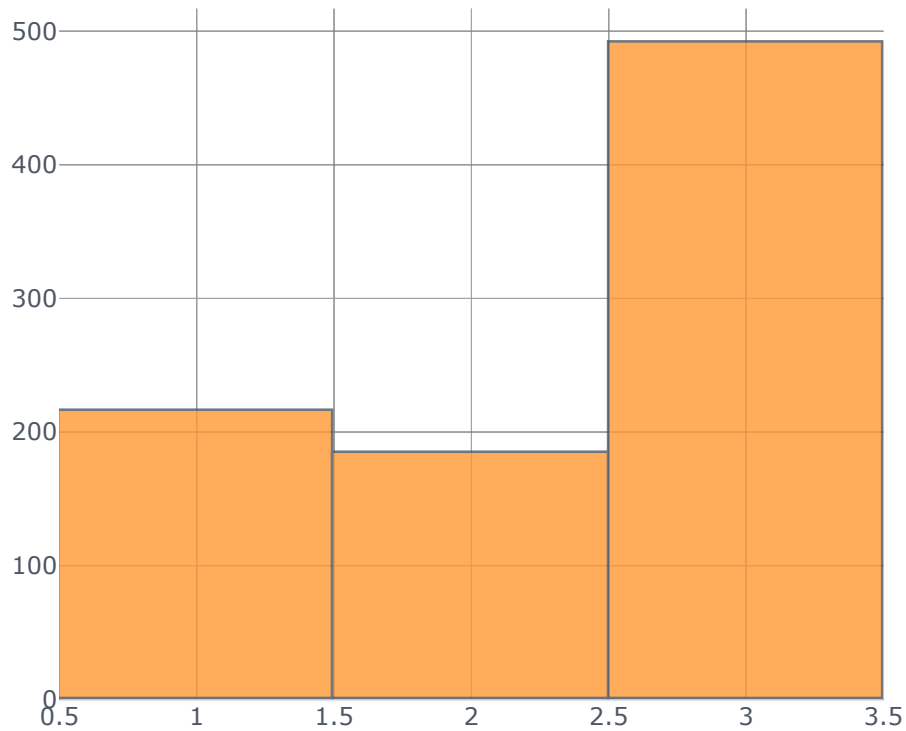


- 나이가 보통 20~40대 가장 많고, 80대 이상도 보인다.

In [15]:



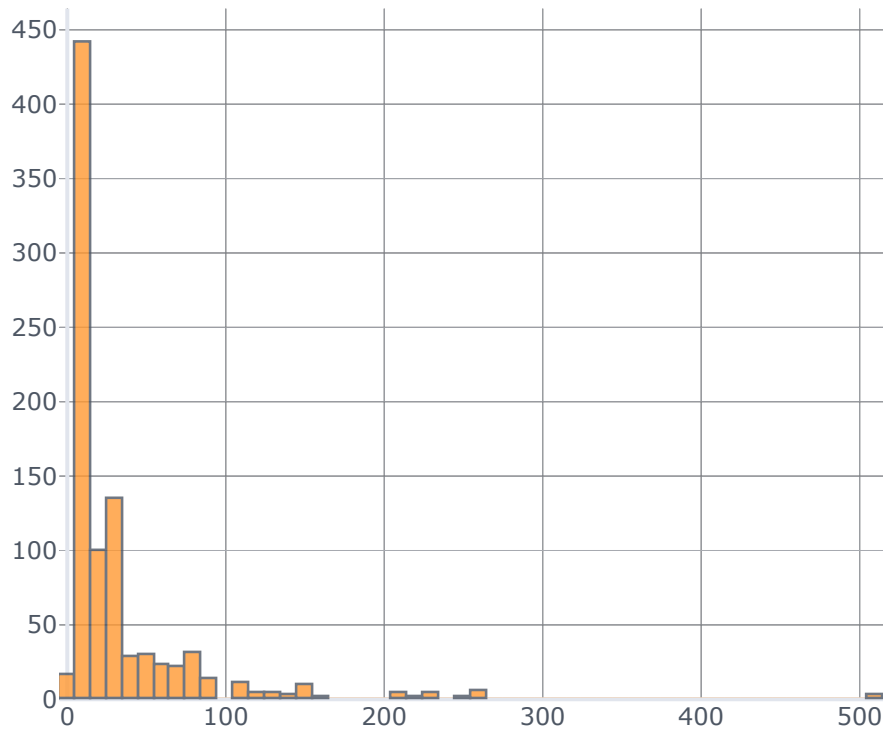
```
train['Pclass'].iplot(kind='hist')
```

[Export to plot.ly »](#)

- Pclass가 3인 사람이 많다.

In [16]:

```
train.Fare.iplot(kind='hist')
```

[Export to plot.ly »](#)

요금이 200, 500인 사람도 일부 있다.

In [17]:

```
train.columns
```

Out[17]:

```
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',  
      'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],  
      dtype='object')
```

In [18]:

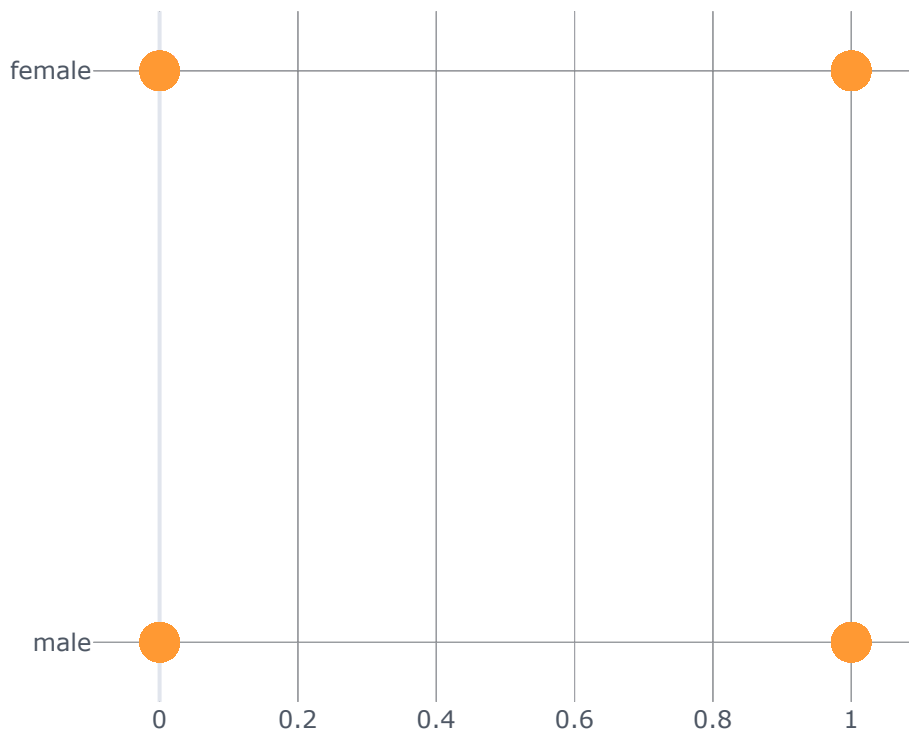
```
train.Ticket
```

Out [18]:

```
0          A/5 21171
1          PC 17599
2  STON/O2. 3101282
3          113803
4          373450
...
886         211536
887         112053
888  W./C. 6607
889         111369
890         370376
Name: Ticket, Length: 891, dtype: object
```

In [19]:

```
train.ipplot(kind='scatter', x='Survived', y='Sex', mode='markers', size=20)
```

[Export to plot.ly »](#)

질문 20~40대 중에 얼마나 생존했나?

REF

- <https://stackabuse.com/using-plotly-library-for-interactive-data-visualization-in-python/>
(<https://stackabuse.com/using-plotly-library-for-interactive-data-visualization-in-python/>)

In []:

